

1/5/1

DIALOG(R) File 351:Derwent WPI  
(c) 2002 Thomson Derwent. All rts. reserv.

A2

010549982 \*\*Image available\*\*

WPI Acc No: 1996-046935/ 199605

XPX Acc No: N96-039426

**Disk array appts for computer system - has data control part which detects abnormality in any buffer memory based on output of detector circuit**

Patent Assignee: OKI ELECTRIC IND CO LTD (OKID )

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
JP 7311658	A	19951128	JP 94128391	A	19940518	199605 B

Priority Applications (No Type Date): JP 94128391 A 19940518

Patent Details:

Patent No	Kind	Lan Pg	Main IPC	Filing Notes
JP 7311658	A	10	G06F-003/06	

Abstract (Basic): JP 7311658 A

The appts (1) includes a number of magnetic disk units (104-108) and a number of buffer memory units (109-113). When data is written from a host computer (2), a data control part distributes the data in each buffer memory unit which stores the data temporarily.

A parity bit from the received data is stored in another buffer memory. The data control part has an abnormal detector circuit (119) which detects any error or abnormality in the contents of the buffer memory unit. The data from the faulty memory is extracted based on the output of the detector circuit.

ADVANTAGE - Improves reliability. Enables holding of data temporarily by other buffer memories.

Dwg.1/14

Title Terms: DISC; ARRAY; APPARATUS; COMPUTER; SYSTEM; DATA; CONTROL; PART; DETECT; ABNORMAL; BUFFER; MEMORY; BASED; OUTPUT; DETECT; CIRCUIT

Derwent Class: T01; T03

International Patent Class (Main): G06F-003/06

International Patent Class (Additional): G06F-011/10; G06F-012/08

File Segment: EPI

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-311658

(43)公開日 平成7年(1995)11月28日

(51)Int.Cl.<sup>6</sup>

G 0 6 F 3/06

11/10

12/08

識別記号

5 4 0

3 3 0 L

3 2 0

庁内整理番号

7608-5B

F I

技術表示箇所

審査請求 未請求 請求項の数4 F D (全 10 頁)

(21)出願番号

特願平6-128391

(22)出願日

平成6年(1994)5月18日

(71)出願人 000000295

沖電気工業株式会社

東京都港区虎ノ門1丁目7番12号

(72)発明者 上村 明利

東京都港区虎ノ門1丁目7番12号 沖電気  
工業株式会社内

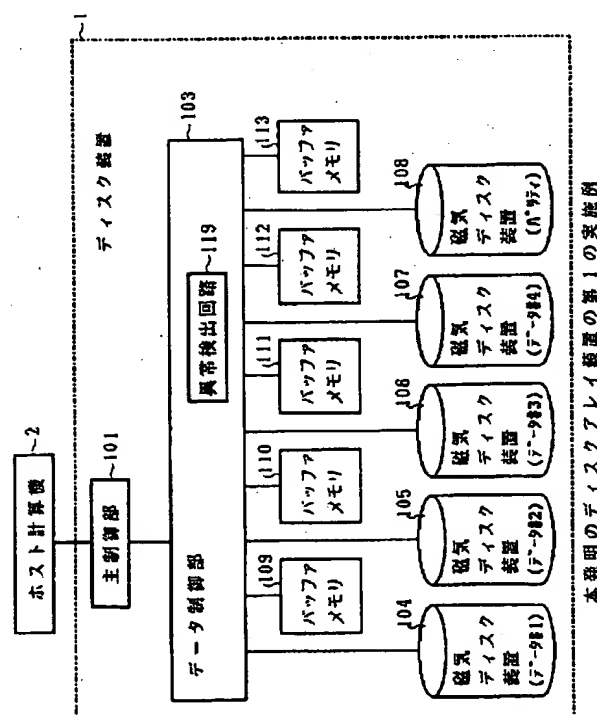
(74)代理人 弁理士 佐藤 幸男

(54)【発明の名称】 ディスクアレイ装置

(57)【要約】

【目的】 バッファメモリの故障に対処可能とし、信頼性を向上させる。

【構成】 ディスクアレイ装置1は、複数の磁気ディスク装置104～108とともに、複数のバッファメモリ109～113を備える。ホスト計算機2からディスクアレイ装置1にデータを書き込む際は、データ制御部103が各バッファメモリ109～113に分散してデータを格納し、そのうちの1つのバッファメモリにはデータから生成したパリティを格納する。データ制御部103は、異常検出回路119によりバッファメモリ109～113のうちのいずれか1つのバッファメモリの異常を検出したとき、他の正常なバッファメモリから故障したバッファメモリに格納されていたデータを生成する。そして、正常な動作を継続する。



## 【特許請求の範囲】

## 【請求項 1】 複数の磁気ディスク装置と、

当該複数の磁気ディスク装置に格納されるデータを一時的に格納する複数のバッファメモリと、

当該各バッファメモリの異常を検出する異常検出回路と、

当該異常検出回路によりいずれか 1 つのバッファメモリが異常であると検出されたとき、当該異常であると検出されたバッファメモリに格納されていたデータを他のバッファメモリから生成して前記磁気ディスク装置に格納するデータ制御部とから成ることを特徴とするディスクアレイ装置。

【請求項 2】 前記各バッファメモリの記憶内容を保持するバッテリと、

ホスト計算機から前記各バッファメモリにデータを書き込んだ時点でライト処理を終了し、前記ホスト計算機によるリード処理及びライト処理とは非同期に前記バッファメモリ中のデータを前記磁気ディスク装置に書き込むよう、前記データ制御部を制御する主制御部とを設けたことを特徴とする請求項 1 記載のディスクアレイ装置。

【請求項 3】 前記主制御部は、前記各バッファメモリのすべてが正常なときは、当該各バッファメモリに前記各バッファメモリにデータを書き込んだ時点でライト処理を終了する一方、前記各バッファメモリのいずれか 1 つ以上が異常なときは、前記各磁気ディスク装置にデータを書き込んだ時点でライト処理を終了するよう、前記データ制御部を制御するようにしたことを特徴とする請求項 2 記載のディスクアレイ装置。

【請求項 4】 前記主制御部は、前記各磁気ディスク装置からデータを読み出す際、前記リード処理で指定された範囲よりも先のデータをも読み出し、前記バッファメモリに格納する一方、前記各磁気ディスク装置からデータを読み出す際、前記各バッファメモリ中にホスト計算機のリード対象であるデータが既に読み出されている場合には、前記磁気ディスク装置にアクセスしないようにしたことを特徴とする請求項 1、請求項 2 又は請求項 3 記載のディスクアレイ装置。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明は、データの信頼性を向上させるとともにアクセス時間を短縮させるようにしたディスクアレイ装置に関するものである。

## 【0002】

【従来の技術】コンピュータシステムに使用されるファイル装置としては、磁気ディスク装置が一般的である。近年、マルチメディア指向に伴ない、取扱うデータ量が増え、ファイルアクセスの高速化が必要になったことから、磁気ディスク装置に置き換えられるものとして、ディスクアレイ装置と呼ばれるファイル装置が注目を浴びている。ディスクアレイ装置とは、複数の磁気ディスク

装置を並列動作させるようにしたもので、カリフォルニア大学バークレー校において考案された RAID と呼ばれるアーキテクチャが適用される。この RAID アーキテクチャにはいくつかのレベルがあるが、その中でレベル 3 を適用したディスクアレイ装置が実用化されている。このようなディスクアレイ装置では、ホスト計算機からのデータを一時的にバッファメモリに格納し、当該バッファメモリ上で分割したデータを各磁気ディスク装置に分散して格納する。

10 【0003】以上が従来の技術の概要であるが、後の実施例と対比するため、従来例の装置の詳細を以下に示す。図 2 は、後述する図 1 の装置と比較される従来の装置を示すブロック図である。図 2 に示す装置は、従来の RAID レベル 3 ディスクアレイ装置である。図 2 において、主制御部 501 は、ディスクアレイ装置 5 の全体を制御する。バッファメモリ 502 は、ランダム・アクセス・メモリ等から成り、データを一時的に格納する。

20 【0004】データ制御部 503 は、データを RAID レベル 3 に合せて処理し、磁気ディスク装置 504 ~ 508 に格納する。磁気ディスク装置 504 ~ 507 は、データをバイト単位に分散して格納する。磁気ディスク装置 508 は、磁気ディスク装置に分散して格納されたデータのパリティを格納する。ホスト計算機 6 は、ディスクアレイ装置 5 にデータのリード/ライト要求を行なう。図 3 は、従来のディスクアレイ装置のデータ及びパリティの分散形式を示す。

30 【0005】ホスト計算機 6 からのデータは、磁気ディスク装置 504 から磁気ディスク装置 507 までに、バイト単位で順に格納される。そして、磁気ディスク装置 504 から磁気ディスク装置 507 までのバイト単位のデータに対して、そのパリティがデータライト時に生成され、磁気ディスク装置 508 に格納される。RAID レベル 3 ディスクアレイ装置では、このようにデータをバイト単位で複数の磁気ディスク装置に分散して格納するため、データを高速にリードライトすることが可能である。また、パリティを格納する磁気ディスク装置 508 を有しているため、この磁気ディスク装置 508 を含む磁気ディスク装置 504 ~ 508 のいずれか 1 台が故障してもディスクアレイ装置としてはデータが失われることがない。これにより、データアクセスの高速性と信頼性が保障される。図 4 は、ディスクアレイ装置のデータ格納単位であるセクタと、内部の磁気ディスク装置のデータ格納単位であるセクタとの対応を示した図である。

50 【0006】磁気ディスク装置 504 ~ 508 のセクタ容量は、現在の磁気ディスク装置の標準である 512 バイトであり、ディスクアレイ装置としてもセクタ容量は磁気ディスク装置の置き換えである 512 バイトとしている。このため、個々の磁気ディスク装置の 1 セクタは

## 3

ディスクアレイ装置の4セクタ分に対応している。例えば、ディスクアレイ装置701の1セクタは、磁気ディスク装置504~507の1セクタの最初の1/4セクタに対応している。図5は、従来のディスクアレイ装置におけるホスト計算機からのデータライト手順のフローチャートである。図2、図3、図4、図5を用いて従来のディスクアレイ装置のデータライトの動作を説明する。

【0007】ホスト計算機6からデータのライト要求が受け付けられると、主制御部501は、まず、磁気ディスク装置504~508の対応セクタを算出する。次に、データを磁気ディスク装置504~508のセクタ単位に合せるために、主制御部501は、データ制御部503を起動して、ステップS360で算出したセクタ単位で磁気ディスク装置504~508よりデータをバッファメモリへリードする(ステップS361)。この時、磁気ディスク装置504~508に分散しているデータは、データ制御部503によりまとめられてバッファメモリ502へ格納される。次にライトすべきデータをホスト計算機6から受け取り、ステップS362でバッファメモリ502へ格納する。ライトデータはステップS361で磁気ディスク装置504~508から読み出したデータの上に上書きするように格納する。最後にバッファメモリ502上のデータ(ライトデータに磁気ディスク装置504~508上のデータを合成してセクタ単位に合せてある)を、データ制御部503を起動して、磁気ディスク装置504~508に分散して格納する。この時、データ制御部503はデータからパリティを生成して磁気ディスク装置508に格納する。

【0008】図6は、従来のディスクアレイ装置におけるホスト計算機からのデータリードのフローチャートである。図2、図3、図4、図5を用いて従来のディスクアレイ装置のデータリードの動作を説明する。

【0009】ホスト計算機6からデータのリード要求が受け付けられると、主制御部501はまず、磁気ディスク装置504~508の対応するセクタを算出する。次に、主制御部501はデータ制御部503を起動して、ステップS370で算出したセクタ単位の先頭からホスト計算機6によって要求されたデータまでの間において、磁気ディスク装置504~508から読み出したデータを捨てる(ステップS371)。次に、ステップS372で、ホスト計算機6によって要求されたデータ間において、データを磁気ディスク装置504~508からデータ制御部503を通して読み出し、主制御部501を通して、ホスト計算機6へ送出する。このとき、磁気ディスク装置504~508に分散しているデータは、データ制御部503によりまとめられて主制御部501へ送られる。最後に、主制御部501はデータ制御部503を起動して、ステップS370で算出したセクタ単位の終了までの間において、磁気ディスク装置50

## 4

4~508から読み出したデータを捨てる。

【0010】次に、磁気ディスク装置504~508のいずれか1台が故障した場合の動作について説明する。この場合、その故障した磁気ディスク装置を使用せずにディスクアレイ装置全体として動作を続けることができる。例えば、前述のステップS372で磁気ディスク装置504から故障が報告された場合を考えると、その時点でデータ制御部503は図示しない内部のパリティ生成回路(本回路はデータライト時にパリティを生成した回路と同じである)により故障した磁気ディスク装置504のデータを、他の磁気ディスク装置505~508のデータにより生成して主制御部501へ送る。

【0011】

【発明が解決しようとする課題】しかしながら、上述した従来の技術には、次のような課題があった。即ち、従来のディスクアレイ装置では、データライト時にライトすべきデータを一時格納しておくバッファメモリが故障した場合、ライトデータが消滅してしまった。この場合、ホスト計算機は、ライト処理を始めからやり直さなければならなかった。従って、システムとしての処理能力が低下するとともに、ディスクアレイ装置としての信頼性が低下するという課題があった。

【0012】

【課題を解決するための手段】本発明のディスクアレイ装置は、上述した課題を解決するため、以下の点を特徴とするものである。

(1)

複数の磁気ディスク装置を備える。

当該複数の磁気ディスク装置に格納されるデータを一時的に格納する複数のバッファメモリを備える。

当該各バッファメモリの異常を検出する異常検出回路を備える。

当該異常検出回路によりいずれか1つのバッファメモリが異常であると検出されたとき、当該異常であると検出されたバッファメモリに格納されていたデータを他のバッファメモリから生成して前記磁気ディスク装置に格納するデータ制御部を備える。

【0013】(2)

(1)において、以下の点を特徴とする。

各バッファメモリの記憶内容を保持するバッテリを備える。

ホスト計算機から各バッファメモリにデータを書き込んだ時点でライト処理を終了するライト制御回路を主制御部内に備える。

ホスト計算機によるリード処理及びライト処理とは非同期に前記バッファメモリ中のデータを磁気ディスク装置に書き込む書き込み制御回路を主制御部内に備える。

【0014】(3)

(2)において、以下の点を特徴とする。

各バッファメモリのすべてが正常なときは、当該各バ

ッファメモリに各バッファメモリにデータを書き込んだ時点でライト処理を終了する一方、各バッファメモリのいずれか1つ以上が異常なときは、各磁気ディスク装置にデータを書き込んだ時点でライト処理を終了するよう、ライト制御回路を切り替える切替制御回路を主制御部に備える。

#### 【0015】(4)

(1)、(2)又は(3)において、以下の点を特徴とする。

各磁気ディスク装置からデータを読み出す際、前記リード処理で指定された範囲よりも先のデータをも読み出し、バッファメモリに格納する読み出し制御回路を主制御部に備える。

各磁気ディスク装置からデータを読み出す際、各バッファメモリ中にホスト計算機のリード対象であるデータが既に読み出されている場合には、磁気ディスクにアクセスしないようにするリード制御回路を主制御部に備える。

#### 【0016】

【作用】データ制御部は、異常検出回路によりいずれか1つのバッファメモリの異常を検出したとき、他の正常なバッファメモリから故障したバッファメモリに格納されていたデータを生成する。この結果、複数のバッファメモリのうちの1つのバッファメモリが故障した場合には、正常な動作を継続することができ、ディスクアレイ装置の信頼性を向上させることができる。また、ホスト計算機からディスクアレイ装置にデータを書き込む際、バッファメモリにデータを書き込んだ時点でホスト計算機に書き込みの終了を通知することにより、磁気ディスク装置の動作を伴わずにデータの書き込みを終了することができる。この結果、ライト動作を高速化することができる。

#### 【0017】

【実施例】以下、本発明の実施例を図面を参照して詳細に説明する。

(第1の実施例) 図1は、本発明のディスクアレイ装置の第1の実施例のブロック図である。この図は、RAIDレベル3-ディスクアレイ装置を示す。図1に示すディスクアレイ装置は、図2に示すディスクアレイ装置と同様に、主制御部101、データ制御部103、複数の磁気ディスク装置104~108を備える。ところが、図2に示すディスクアレイ装置とは異なり、主制御部101とデータ制御部103の間にバッファメモリは設けられておらず、データ制御部103と接続された複数のバッファメモリ109~113が設けられている。これに伴ない、データ制御部103の制御手順は、図2のデータ制御部503の制御手順とは多少異なる。データ及びパリティ分散形式、セクタの対応形式は、図3及び図4に示す従来のものと同様である。図7は、本発明のディスクアレイ装置におけるホスト計算機からのデータラ

イトのフローチャートである。図1、図7を用いて本装置のデータライトの動作を説明する。

【0018】ホスト計算機2からデータのライト要求が受け付けられると、主制御部101は、まず、磁気ディスク装置104~108の対応セクタを算出する。次に、データを磁気ディスク装置104~108のセクタ単位に合せるために、主制御部101は、データ制御部103を起動して、ステップS301で算出したセクタ単位で磁気ディスク装置104~108よりデータ及びパリティをバッファメモリ109~113へリードする(ステップS302)。この時、磁気ディスク装置104~108に分散しているデータ及びパリティは、そのまま各バッファメモリ109~113へ格納される。次にライトすべきデータをホスト計算機2から受け取り、データ制御部103を起動してバッファメモリ109~113へ格納する。この時、データ制御部103はステップS303でデータからパリティを生成してバッファメモリ113へ格納する。ライトデータはステップS302で磁気ディスク装置104~108から読み出したデータの上に上書きするように格納する。最後にバッファメモリ109~113上のデータ(ライトデータに磁気ディスク装置104~108上のデータを合成してセクタ単位に合せてある)を、データ制御部を起動して、磁気ディスク装置104~108に分散して格納する(ステップS304)。リード動作については、従来と同じであるので重複する説明を省略する。

【0019】次に、バッファメモリの異常検出回路119がバッファメモリの異常を検出した場合の動作について説明する。異常検出回路119は、バッファメモリ109~113のパリティエラーを検出することにより、バッファメモリ109~113のいずれかの異常を検出する。この場合、異常となったバッファメモリ109~113は使用せずにディスクアレイ装置全体として動作を続けることができる。例えば、前述のステップS304でバッファメモリ109が異常であることが検出された場合を考えると、その時点でデータ制御部103はデータフローを図8のように切り替える。図8で、バッファメモリ109中のデータは使用不可能であるが、データ制御部103は内部のパリティ生成回路(本回路はデータライト時にパリティを生成する回路と同じである)120により故障したバッファメモリ109のデータを、他のバッファメモリ110~113のデータにより生成して主制御部101へ送る。

【0020】このようにしてバッファメモリ109~113のいずれか1台が故障してもディスクアレイ装置全体としては動作を継続できるので、高速性及び信頼性を保障することができる。また、従来と同様に、磁気ディスク装置104~108のいずれか1台が同時に故障した場合も動作を継続することができる。この場合、バッファメモリ109~113と磁気ディスク装置104~

108はいずれの1台ずつが故障しても動作を継続することができる。

【0021】(第2の実施例)図9は、本発明の第2の実施例に係るRAIDレベル3-ディスクアレイ装置の構成図である。図9においては、各バッファメモリ109~113にバッテリー114~118がそれぞれ備えられている。これらのバッテリー114~118は、各バッファメモリ109~113の記憶内容を保持する。図10、図11は、本発明のディスクアレイ装置におけるホスト計算機からのデータライト手順のフローチャートである。

【0022】図9、図10、図11を用いて本装置のデータライトの動作を説明する。ホスト計算機2からデータのライト要求が受け付けられると、主制御部101はライトすべきデータをホスト計算機2から受け取り、データ制御部103を起動してバッファメモリ109~113へ格納する(ステップS310)。この時、データ制御部103はデータからパリティを生成してバッファメモリ113へ格納する。主制御部101はこの時点でホスト計算機2へライト処理の終了を通知する。このようにして磁気ディスク装置104~108の動作を伴わずにライト動作を終了させることができるので、ライト動作の大幅な高速化を図ることができる。

【0023】バッファメモリ109~113へ格納したデータ及びパリティは、ホスト計算機2からのリードライト要求が来ない間に、図11のフローチャートに従い、磁気ディスク装置104~108に格納される。格納の開始時は、まず、ステップS320で磁気ディスク装置104~108の対応セクタを算出する。次に、データを磁気ディスク装置104~108のセクタ単位に合せるために、主制御部101は、データ制御部103を起動して、ステップS320で算出したセクタ単位で磁気ディスク装置よりデータ及びパリティをバッファメモリ109~113へリードする(ステップS321)。この時、磁気ディスク装置104~108に分散して格納されているデータ及びパリティは、そのまま各バッファメモリ109~113へ格納するが、リードしたデータ及びパリティの格納は、ライトデータの範囲外の部分のみ行なう。最後にバッファメモリ109~113上のデータ及びパリティ(ライトデータに磁気ディスク装置上のデータを合成してセクタ単位に合せてある)を、データ制御部103を起動して、対応する磁気ディスク装置104~108に格納する。

【0024】リード動作については、従来と同じであるので重複する説明を省略する。尚、本実施例において、バッファメモリ109~113は、バッテリー114~118によりバックアップされているので、バッファメモリ109~113中にライトデータを保持した状態で本装置の電源が落されても、ライトデータは正常に保持できる。

【0025】(第3の実施例)第3の実施例におけるディスクアレイ装置のハードウェア構成は、第2の実施例と同じである。図12は、第3の実施例のディスクアレイ装置におけるホスト計算機からのデータライト手順のフローチャートである。以下、図9において、図12を参照しつつ、第3の実施例のディスクアレイ装置の動作を説明する。

【0026】ホスト計算機2からデータのライト要求が受け付けられると、主制御部101は、まず、ステップS330でバッファメモリ109~113の状態をチェックする。バッファメモリ109~113がすべて正常であれば、第2の実施例と同様にライトデータをバッファメモリ109~113に格納して(ステップS331)、ライト処理を終了する。ステップS331の詳細は、図10と同様である。バッファメモリ109~113に格納したデータ及びパリティは、ホスト計算機2からのリードライト要求が来ない間に、第2の実施例と同様に図11のフローチャートに従い、磁気ディスク装置104~108に格納される。一方、ステップS330でいずれか1台のバッファメモリが正常でなければ、正常な残りのバッファメモリを利用してデータを磁気ディスク装置104~108に格納してからライトを終了する(ステップS332)。ステップS332の詳細は図7と同様である。

【0027】このようにしてすべてのバッファメモリ109~113が正常なときのみ第2の実施例と同様の高速化ライトが行なわれ、1台のバッファメモリが故障し、バッファメモリ全体としてデータの信頼性が低下しているときは、バッファメモリにライトデータを残さないようにしたので、第2の実施例の効果に加えて信頼性の向上を図ることができる。リード動作については従来と同じであるので省略する。尚、本実施例においても、第2の実施例と同様に、バッファメモリは、バッテリーによりバックアップされているので、バッファメモリ中にライトデータを保持した状態で本装置の電源が落されても、ライトデータは正常に保持できる。

【0028】(第4の実施例)第4の実施例に係るディスクアレイ装置のハードウェア構成は、第1の実施例と同じである。図13は、第4の実施例のディスクアレイ装置におけるホスト計算機からのデータライト手順のフローチャートである。以下、図1において、図13を参照しつつ、第4の実施例に係る装置のデータライトの動作を説明する。

【0029】ホスト計算機2からデータのライト要求が受け付けられると、ステップS340で主制御部101はまず磁気ディスク装置104~108の対応セクタを算出する。次に、算出したセクタ単位のデータがバッファメモリ109~113中に存在しているかをステップS341で調べる。データが存在していれば、ライトすべきデータをホスト計算機2から受け取り、データ制御

部103を起動して、バッファメモリ109~113へ格納する(ステップS343)。ライトデータは、磁気ディスク装置104~108から読み出してあるデータの上に上書きするようにして格納する。最後に、バッファメモリ109~113上のデータ(ライトデータに磁気ディスク装置104~108上のデータを合成してセクタ単位に合せてある)をデータ制御部103を起動して、対応する磁気ディスク装置104~108に格納する。

【0030】一方、ステップS341で、データが存在していなければ、データを磁気ディスク装置104~108のセクタ単位に合せるために、主制御部101はデータ制御部103を起動して、ステップS340で算出したセクタ単位で磁気ディスク装置104~108からデータ及びパリティをバッファメモリ109~113にリードする(ステップS342)。この時、磁気ディスク装置104~108に分散しているデータ及びパリティは、そのまま対応するバッファメモリ109~113に格納される。尚、算出した範囲より先のセクタのデータもこのときいっしょに読み出し、バッファメモリ109~113にリードする。

【0031】図14は、第4の実施例に係るディスクアレイ装置におけるホスト計算機からのデータリード手順のフローチャートである。以下、図9において、図14を参照しつつ、第4の実施例の装置のデータリードの動作を説明する。

【0032】ホスト計算機2からデータのリード要求が受け付けられると、ステップS350で主制御部101はまず、磁気ディスク装置104~108の対応セクタを算出する。次に、算出したセクタ単位のデータが既にバッファメモリ109~113中に存在しているかをステップS351で調べる。データが存在していれば、リードすべきデータを、データ制御部103を起動して、バッファメモリ109~113からホスト計算機2に送出する(ステップS352)。この時、バッファメモリ109~113に分散して格納されているデータは、データ制御部103によりまとめられて主制御部101へ送られる。ステップS351で、データが存在していなければ、主制御部101はデータ制御部103を起動して、ステップS350で算出したセクタ単位の先頭からホスト計算機2によって要求されたデータまでの間において、磁気ディスク装置104~108から読み出したデータを捨てる(ステップS353)。

【0033】次に、ホスト計算機2によって要求されたデータ間において、データを磁気ディスク装置104~108からデータ制御部103を通して読み出し、主制御部101を通して、ホスト計算機2へ送出する(ステップS354)。この時、磁気ディスク装置に分散しているデータは、データ制御部103によりまとめられて主制御部101に送られる。最後に、主制御部101

は、データ制御部103を起動して、ステップS350で算出したセクタ単位の終了までの間において、磁気ディスク装置104~108から読み出したデータを捨てる(ステップS355)。尚、算出した範囲より先のセクタのデータもこのときいっしょに読み出し、バッファメモリ109~113にリードする。

【0034】このようにして磁気ディスク装置104~108からデータを読み出す際、指定された範囲よりも先のデータも読み出されるので、磁気ディスク装置104~108をアクセスせずにバッファメモリ109~113をアクセスしてリード処理を済ませることができる場合を多くし、リード動作の高速化を図ることができる。尚、本発明は上述した実施例に限定されるものではなく、種々の変形が可能であることはもちろんである。例えば、磁気ディスク装置104~108の数と、バッファメモリ109~113の数を同数としているが、数を揃える必然性はない。

#### 【0035】

【発明の効果】以上説明したように、本発明のディスクアレイ装置によれば、複数のバッファメモリを設けるとともに、当該各バッファメモリの異常を検出する異常検出回路を設けるようにしたので、1台までのバッファメモリの故障に対しては残りのバッファメモリによりデータを正常に保持することができ、ディスクアレイ装置としての信頼性を向上させることができる。

#### 【図面の簡単な説明】

【図1】本発明のディスクアレイ装置の第1の実施例のブロック図である。

【図2】従来のディスクアレイ装置の構成を示すブロック図である。

【図3】従来のディスクアレイ装置のデータとパリティの分散手順を示すブロック図である。

【図4】従来のディスクアレイ装置のセクタの対応関係の説明図である。

【図5】従来のデータライトの手順を説明するフローチャートである。

【図6】従来のデータリードの手順を説明するフローチャートである。

【図7】本発明のデータライトの手順を説明するフローチャートである。

【図8】本発明のディスクアレイ装置におけるデータフローの説明図である。

【図9】本発明のディスクアレイ装置の第2の実施例のブロック図である。

【図10】本発明の第2の実施例のデータライト手順を説明するフローチャートである。

【図11】本発明の第2の実施例の磁気ディスク装置へのデータライト手順を説明するフローチャートである。

【図12】本発明の第3の実施例のデータライト手順を説明するフローチャートである。

【図13】本発明の第4の実施例のデータライト手順を説明するフローチャートである。

【図14】本発明の第4の実施例のデータリード手順を説明するフローチャートである。

【符号の説明】

1 ディスクアレイ装置

2 ホスト計算機

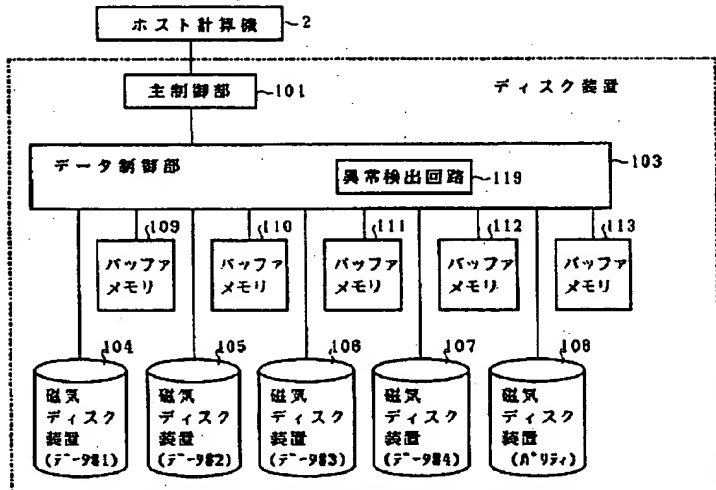
101 主制御部

103 データ制御部

104～108 磁気ディスク装置

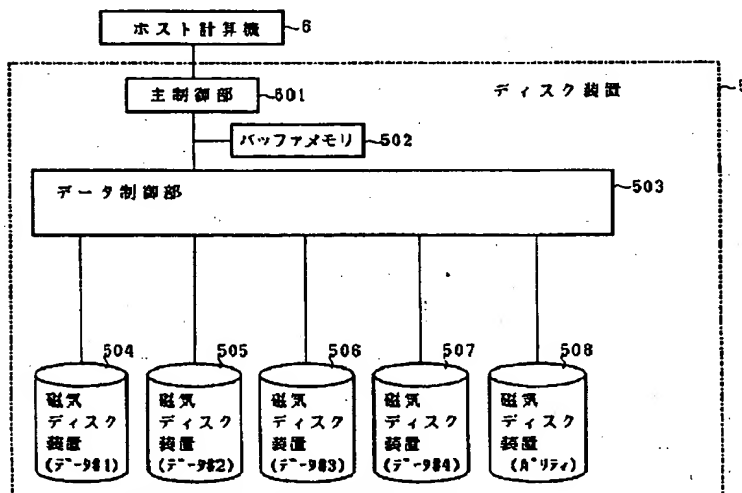
109～113 バッファメモリ

【図1】



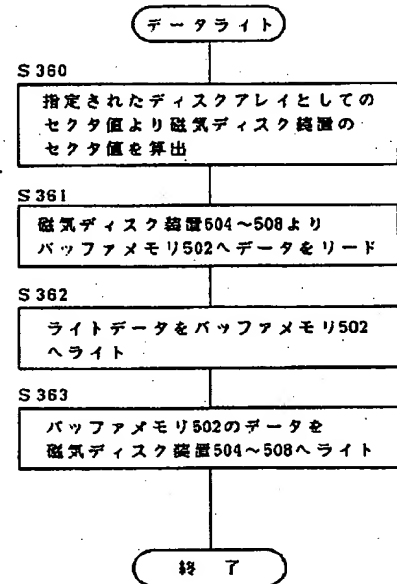
本発明のディスクアレイ装置の第1の実施例

【図2】



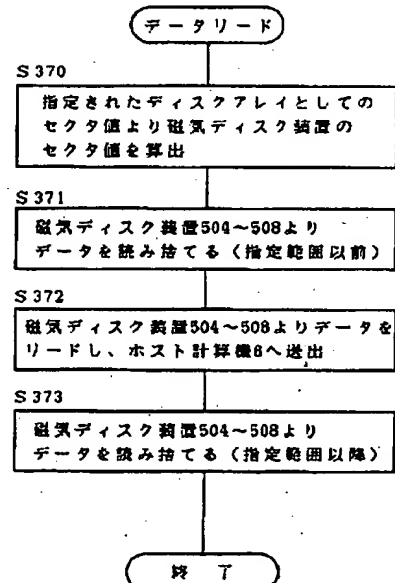
従来のディスクアレイ装置の一例

【図5】



従来のデータライト手順

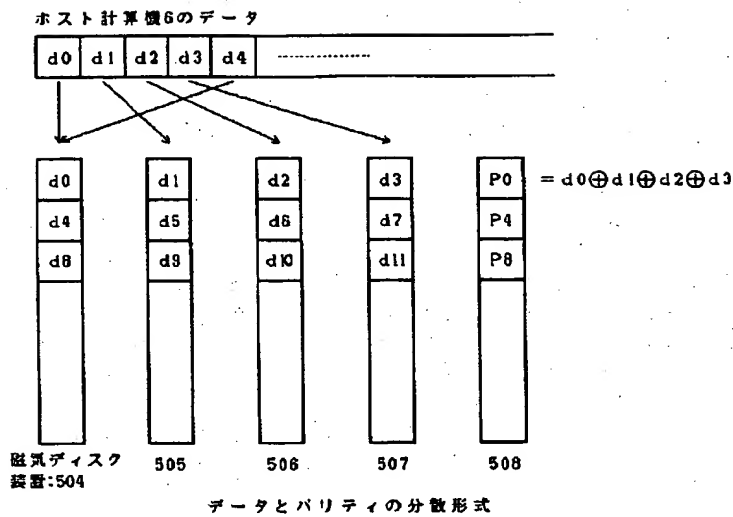
【図6】



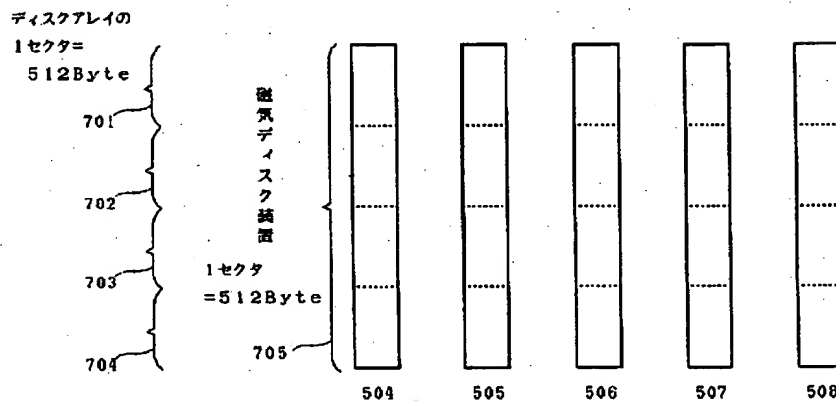
従来のデータリード手順



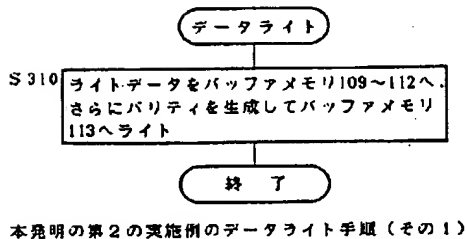
【図3】



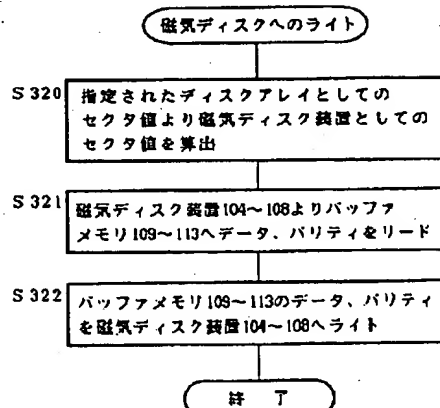
【図4】



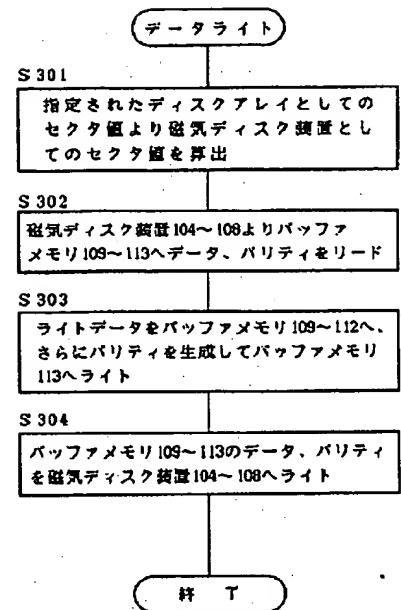
【図10】



【図11】

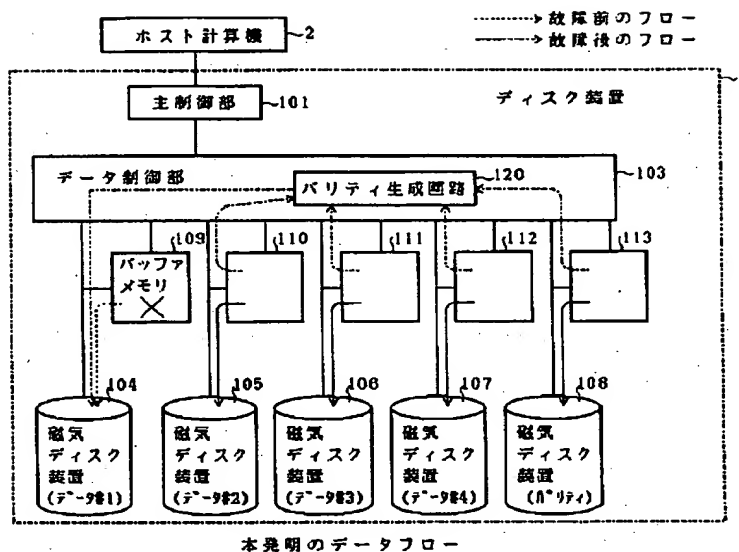


【図7】

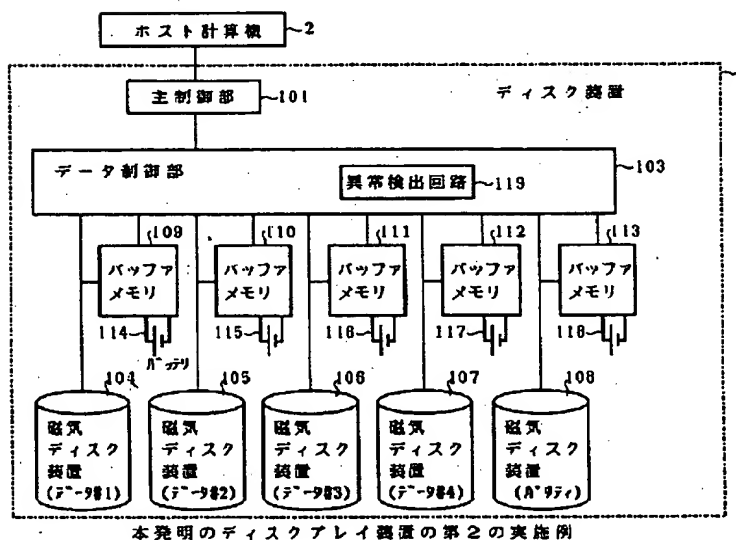


本発明の第1の実施例のデータライト手順

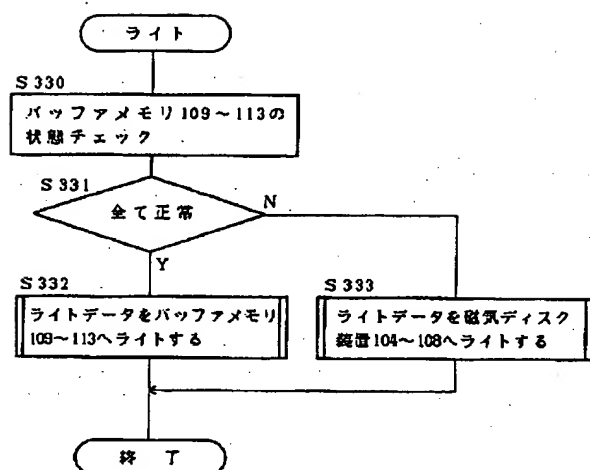
【図8】



【図9】

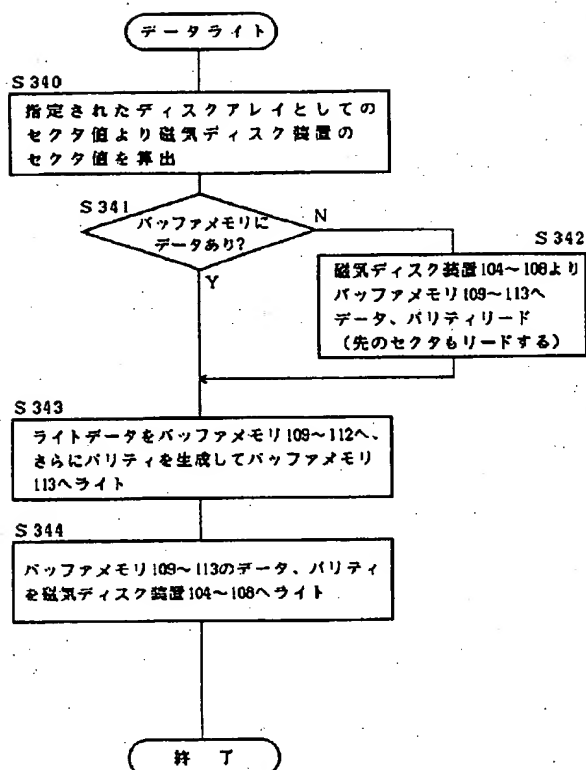


【図12】



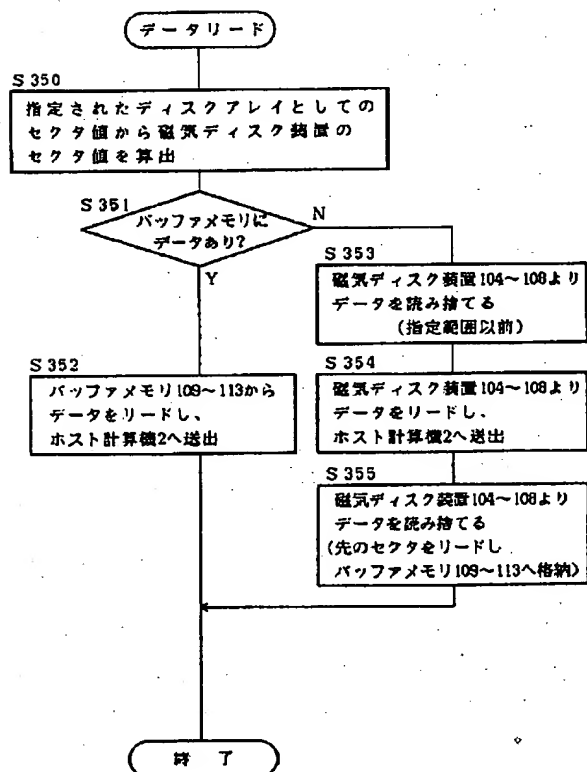
本発明の第3の実施例のデータライト手順

【図13】



本発明の第4の実施例のデータライト手順

【図14】



本発明の第4の実施例のデータリード手順